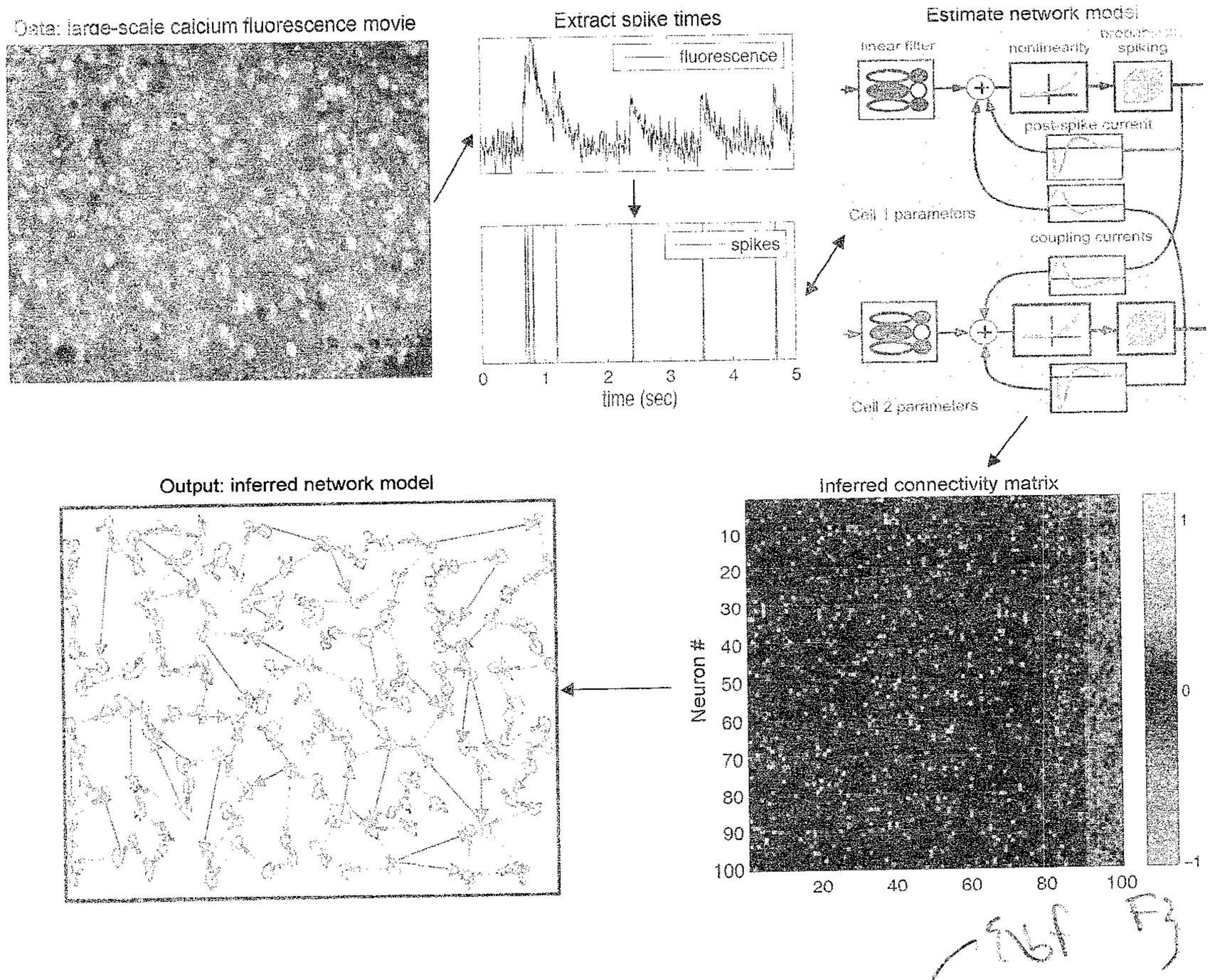


FIG 1. *Schematic overview. The raw observed data is a large-scale calcium fluorescence movie, which is preprocessed to correct for movement artifacts and find regions-of-interest, i.e., putative neurons. (Note that we have omitted details of these important preprocessing steps in this paper, see, e.g., (Cossart et al., 2003; Dombeck et al., 2007) for further details.) Given the fluorescence traces $F_i(t)$ from each neuron, we estimate the underlying spike trains (i.e., the time series of neural activity) using statistical deconvolution methods. Then we estimate the parameters of a network model given the observed data. Our major goal is to obtain an accurate estimate of the network connectivity matrix, which summarizes the information we are able to infer about the local neuronal microcircuit. (We emphasize that this illustration is strictly schematic, and does not correspond directly to any of the results described below.) This figure adapted from personal communications with R. Yuste, B. Watson, and A. Packer.*

## 2. Methods.

2.1. *Model.* We begin by detailing a parametric generative model for the (unobserved) joint spike trains of all $N$ observable neurons, along with the observed calcium fluorescence data. Each neuron is modeled as a generalized linear model (GLM). This class of models is known to capture the statistical firing properties of individual neurons fairly accurately (Brillinger, 1988; Chornoboy et al., 1988; Brillinger, 1992; Plesser and Gerstner, 2000; Paninski et al., 2004; Paninski, 2004; Rigat et al., 2006; Truccolo et al., 2005; Nykamp, 2007; Kulkarni and Paninski, 2007; Pillow et al., 2008; Vidne et al., 2009; Stevenson et al., 2009). We denote the $i$-th neuron's activity at time $t$ as $n_i(t)$: in continuous time, $n_i(t)$ could be modeled as an unmarked point process, but we will take a discrete-time approach here, with each $n_i(t)$

that this model generalizes (via a simple augmentation of the state variable $h_{ij}(t)$) to allow each neuron pair to have several spike history terms, each with a unique time constant, which when weighted and summed allow us to model a wide variety of possible post-synaptic effects, including bursting, facilitating, and depressing synapses; see (Vogelstein et al., 2009) for further details. We restrict our attention to the case of a single time constant $\tau_{ij}^h$ per synapse here, so the deterministic part of $h_{ij}(t)$ is a simple exponentially-filtered version of the spike train $n_j(t)$. Furthermore, we assume that $\tau_{ij}^h$ is the same for all neurons and all synapses, although in principle each synapse could be modeled with its unique $\tau_{ij}^h$. We do that both for simplicity and also because we find that the detailed shape of the coupling terms $h_{ij}(t)$ had a limited effect on the inference of the connectivity matrix, as illustrated in Figure 12 below. Thus, we treat $\tau_{ij}^h$ and $\sigma_{ij}^h$ as known synaptic parameters which are the same for each neuron pair $(i, j)$, and denote them as $\tau_h$ and $\sigma_h$ hereafter. We chose values for $\tau_h$ and $\sigma_h$ in our inference based on experimental data (Lefort et al., 2009); see Table 1 below. Therefore our unknown spiking parameters are $\{\mathbf{w}_i, k_i, b_i\}_{i \le N}$, with $\mathbf{w}_i = (w_{i1}, \ldots, w_{iN})$.

The problem of estimating the connectivity parameters $\mathbf{w} = \{\mathbf{w}_i\}_{i \le N}$ in this type of GLM, given a fully-observed ensemble of neural spike trains $\{n_i(t)\}_{i \le N}$, has recently received a great deal of attention; see the references above for a partial list. In the calcium fluorescent imaging setting, however, we do not directly observe spike trains; $\{n_i(t)\}_{i \le N}$ must be considered a hidden variable here. Instead, each spike in a given neuron leads to a rapid increase in the intracellular calcium concentration, which then decays slowly due to various cellular buffering and extrusion mechanisms. We in turn make only noisy, indirect, and subsampled observations of this intracellular calcium concentration, via fluorescent imaging techniques (Yuste et al., 2006). To perform statistical inference in this setting, (Vogelstein et al., 2009) proposed a simple conditional first-order hidden Markov model (HMM) for the intracellular calcium concentration $C_i(t)$ in cell $i$ at time $t$, along with the observed fluorescence, $F_i(t)$:

$$C_i(t) = C_i(t - \Delta) + \left(C_i^b - C_i(t - \Delta)\right) \Delta / \tau_i^c + A_i n_i(t) + \sigma_i^c \sqrt{\Delta} \epsilon_i^c(t), \tag{4}$$

$$F_i(t) = \alpha_i S(C_i(t)) + \beta_i + \sqrt{(\sigma_i^F)^2 + \gamma_i S(C_i(t))} \epsilon_i^F(t). \tag{5}$$

This model can be interpreted as a simple driven autoregressive process: under nonspiking conditions, $C_i(t)$ fluctuates around the baseline level of $C_i^b$, driven by normally-distributed noise $\epsilon_i^c(t)$ with standard deviation $\sigma_i^c \sqrt{\Delta}$. Whenever the neuron fires a spike, $n_i(t) = 1$, the calcium variable $C_i(t)$ jumps by a fixed amount $A_i$, and subsequently decays with time constant $\tau_i^c$. The fluorescence signal $F_i(t)$ corresponds to the count of photons collected at the detector per neuron per imaging frame. This photon count may be modeled with normal statistics, with the mean given by a saturating Hill-type function $S(C) = C/(C + K_d)$ (Yasuda et al., 2004) and the variance scaling with the mean; see (Vogelstein et al., 2009) for further discussion. Because the parameter $K_d$ effectively acts as a simple scale factor, and is a property of the fluorescent indicator, we assume throughout this work that it is known. Figure 3 shows a couple examples depicting the relationship between spike trains and observations. It will be useful to define an effective SNR as

$$\text{eSNR} = \frac{E[F_i(t) - F_i(t - \Delta) \mid n_i(t) = 1]}{E[(F_i(t) - F_i(t - \Delta))^2/2 \mid n_i(t) = 0]^{1/2}}, \tag{6}$$

i.e., the size of a spike-driven fluorescence jump divided by a rough measure of the standard deviation of the baseline fluorescence. For concreteness, the effective SNR values in Fig. 3 were 9 and 3 in the left and right panels, respectively.

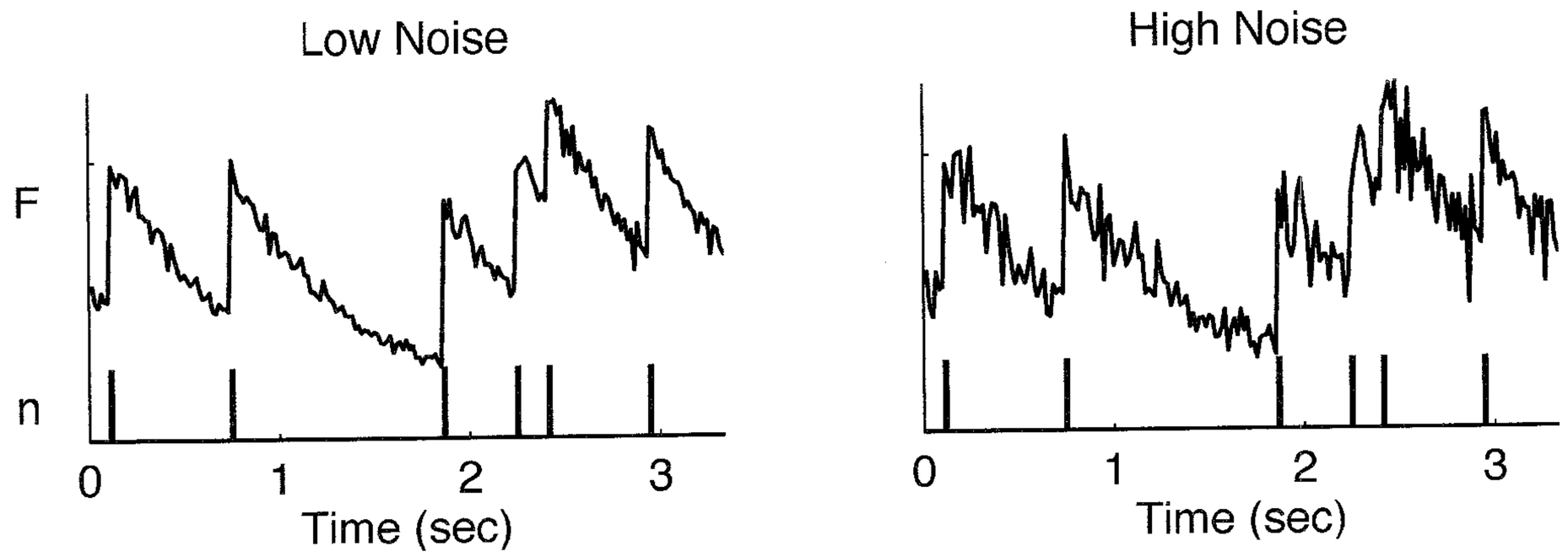


FIG 3. *Two example traces of simulated fluorescence data, at different SNR levels, demonstrating the relationship between spike trains and observed fluorescence in our model. Note that both panels have the same underlying spike train. Simulation parameters:* $k_i = 0.7$, $C_i^b = 1\ \mu M$, $\tau_i^c = 500\ msec$, $A_i = 50\ \mu M$, $\sigma_i^c = 0.1\ \mu M$. $\gamma_i = 0.004$ *(effective SNR* $\approx 9$*, as defined in Eq. (6); see also Figure 9 below) in the left panel and* $\gamma_i = 0.016$ *(eSNR* $\approx 3$*) in the right panel, and* $\sigma_i^F = 0$, $\Delta = (60\ Hz)^{-1}$.

To summarize, Eqs. (1-5) define a coupled HMM: the underlying spike trains $\{n_i(t)\}_{i \le N}$ and spike history terms $\{h_{ij}(t)\}_{i,j \le N}$ evolve in a Markovian manner given the stimulus $S^{ext}(t)$. These spike trains in turn drive the intracellular calcium concentrations $\{C_i(t)\}_{i \le N}$, which are themselves Markovian, but evolving at a slower timescale $\tau_i^c$. Finally, we observe only the fluorescence signals $\{F_i(t)\}_{i \le N}$, which are related in a simple Markovian fashion to the calcium variables $\{C_i(t)\}_{i \le N}$.

2.2. *Goal and general strategy.* Our primary goal is to estimate the connectivity matrix, $\mathbf{w}$, given the observed set of calcium fluorescence signals $\mathbf{F} = \{\mathbf{F}_i\}_{i \le N}$, where $\mathbf{F}_i = \{F_i(t)\}_{t \le T}$. We must also deal with a number of intrinsic parameters[2], $\theta_i$: the intrinsic spiking parameters[3] $\{b_i, w_{ii}\}_{i \le N}$, the calcium parameters $\{C_i^b, \tau_i^c, A_i, \sigma_i^c\}_{i \le N}$, and the observation parameters $\{\alpha_i, \beta_i, \gamma_i, \sigma_i^F\}_{i \le N}$. We addressed the problem of estimating these intrinsic parameters in earlier work (Vogelstein et al., 2009); thus our focus here will be on the connectivity matrix $\mathbf{w}$. A Bayesian approach is natural here, since we have a good deal of prior information about neural connectivity; see (Rigat et al., 2006) for a related discussion. However, a fully-Bayesian approach, in which we numerically integrate over the very high-dimensional parameter space $\theta = \{\theta_i\}_{i \le N}$, where $\theta_i = \{\mathbf{w}_i, b_i, C_i^b, \tau_i^c, A_i, \sigma_i^c, \alpha_i, \beta_i, \gamma_i, \sigma_i^F\}$, is less attractive from a computational point of view. Thus, our compromise is to compute *maximum a posteriori* (MAP) estimates for the parameters via an expectation-maximization (EM) algorithm, in which the sufficient statistics are computed by a hybrid blockwise Gibbs sampler and sequential Monte

[2]The intrinsic parameters for neuron $i$ are all its parameters, minus the cross-coupling terms, i.e. $\tilde{\theta}_i = \theta_i \backslash \{w_{ij}\}_{i \neq j}$.

[3]To reduce the notational load, we will ignore the estimation of the stimulus filter $k_i$ below; this term may be estimated with $b_i$ and $w_{ii}$ using very similar convex optimization methods, as discussed in (Vogelstein et al., 2009).

Carlo (SMC) method. More specifically, we iterate the steps:

$$\textbf{E step: Evaluate } Q(\theta,\theta^{(l)}) = E_{P[\mathbf{X}|\mathbf{F};\theta^{(l)}]} \ln P[\mathbf{F},\mathbf{X}|\theta] = \int P[\mathbf{X}|\mathbf{F};\theta^{(l)}] \ln P[\mathbf{F},\mathbf{X}|\theta] d\mathbf{X}$$

$$\textbf{M step: Solve } \theta^{(l+1)} = \operatorname*{argmax}_{\theta} \left\{ Q(\theta,\theta^{(l)}) + \ln P(\theta) \right\},$$

where $\mathbf{X}$ denotes the set of all hidden variables $\{C_i(t), n_i(t), h_{ij}(t)\}_{i,j\le N, t\le T}$ and $P(\theta)$ denotes a (possibly improper) prior on the parameter space $\theta$. According to standard EM theory (Dempster et al., 1977; McLachlan and Krishnan, 1996), each iteration of these two steps is guaranteed to increase the log-posterior $\ln P(\theta^{(l)}|\mathbf{F})$, and will therefore lead to at least a locally maximum a posteriori estimator.

Now, our major challenge is to evaluate the auxiliary function $Q(\theta,\theta^{(l)})$ in the E-step. Our model is a coupled HMM, as discussed in the previous section; therefore, as usual in the HMM setting (Rabiner, 1989), $Q$ may be broken up into a sum of simpler terms:

$$\begin{aligned} Q(\theta,\theta^{(l)}) &= \sum_{it} \int \ln P[F_i(t)|C_i(t);\alpha_i,\beta_i,\gamma_i,\sigma_i^F] dP[C_i(t)|\mathbf{F};\theta^{(l)}] \\ &+ \sum_{it} \int \ln P[C_i(t)|C_i(t-\Delta),n_i(t);C_i^b,\tau_i^c,A_i,\sigma_i^c] dP[C_i(t),C_i(t-\Delta)|\mathbf{F};\theta^{(l)}] \\ &+ \sum_{it} \int \ln P[n_i(t)|\mathbf{h}_i(t);b_i,\mathbf{w}_i] dP[n_i(t),\mathbf{h}_i(t)|\mathbf{F};\theta^{(l)}], \end{aligned} \tag{7}$$

where $\mathbf{h}_i(t) = \{h_{ij}(t)\}_{j\le N}$. Note that each of the three sums here corresponds to a different component of the model described in Eqs. (1-5): the first sum involves the fluorescent observation parameters, the second the calcium dynamics, and the third the spiking dynamics.

Thus we need only compute low-dimensional marginals of the full posterior distribution $P[\mathbf{X}|\mathbf{F};\theta]$; specifically, we need the pairwise marginals $P[C_i(t)|\mathbf{F};\theta]$, $P[C_i(t),C_i(t-\Delta)|\mathbf{F};\theta]$, and $P[n_i(t),\mathbf{h}_i(t)|\mathbf{F};\theta]$. Details for calculating $P[C_i(t),C_i(t-\Delta)|\mathbf{F}_i;\theta_i]$ and $P[C_i(t)|\mathbf{F}_i;\theta_i]$ are found in (Vogelstein et al., 2009), while calculating the joint marginal for the high dimensional hidden variable $\mathbf{h}_i$ necessitates the development of specialized blockwise Gibbs-SMC sampling methods, as we describe in the subsequent sections 2.3 and 2.4. Once we have obtained these marginals, the M-step breaks up into a number of independent optimizations that may be computed in parallel and which are therefore relatively straightforward (Section 2.5); see section 2.6 for a pseudocode summary along with some specific implementation details.

2.3. *Initialization of intrinsic parameters via sequential Monte Carlo methods.* We begin by constructing relatively cheap, approximate preliminary estimators for the intrinsic parameters, $\tilde{\theta}_i$. The idea is to initialize our estimator by assuming that each neuron is observed independently. Thus we want to compute $P[C_i(t),C_i(t-\Delta)|\mathbf{F}_i;\tilde{\theta}_i]$ and $P[C_i(t)|\mathbf{F}_i;\tilde{\theta}_i]$, and solve the M-step for each $\tilde{\theta}_i$, with the connectivity matrix parameters held fixed. This single-neuron case is much simpler, and has been discussed at length in (Vogelstein et al., 2009); therefore, we only provide a brief overview here. The standard forward and backward recursions provide the necessary posterior distributions, in principle (Shumway and Stoffer,

for more details, see (Vogelstein et al., 2009). Thus equations (10-14) may be used to compute the sufficient statistics for estimating the intrinsic parameters $\tilde{\theta}_i$ for each neuron.

As discussed following Eq. (7), the M-step decouples into three independent subproblems. The first term depends on only $\{\alpha_i, \beta_i, \gamma_i, \sigma_i\}$; since $P[F_i(t)|S(C_i(t)); \tilde{\theta}_i]$ is Gaussian, we can estimate these parameters by solving a weighted regression problem (specifically, we use a coordinate-optimization approach: we solve a quadratic problem for $\{\alpha_i, \beta_i\}$ while holding $\{\gamma_i, \sigma_i\}$ fixed, then estimate $\{\gamma_i, \sigma_i\}$ by the usual residual error formulas while holding $\{\alpha_i, \beta_i\}$ fixed). Similarly, the second term requires us to optimize over $\{\tau_i^c, A_i, C_i^b\}$, and then we use the residuals to estimate $\sigma_i^c$. Note that all the parameters mentioned so far are constrained to be non-negative, but may be solved efficiently using standard quadratic program solvers if we use the simple reparameterization $\tau_i^c \to 1 - \Delta/\tau_i^c$. Finally, the last term may be expanded:

$$(15)\quad E[\ln P[n_i(t), \mathbf{h}_i(t)|\mathbf{F}; \theta_i]] = P[n_i(t), \mathbf{h}_i(t)|\mathbf{F}; \theta_i] \ln f[J_i(t)] + (1 - P[n_i(t), \mathbf{h}_i(t)|\mathbf{F}; \theta_i]) \ln[1 - f(J_i(t))];$$

since $J_i(t)$ is a linear function of $\{b_i, \mathbf{w}_i\}$, and the right-hand side of Eq. (15) is concave in $J_i(t)$, we see that the third term in Eq. (7) is a sum of terms which are concave in $\{b_i, \mathbf{w}_i\}$ — and therefore also concave in the linear subspace $\{b_i, w_{ii}\}$ with $\{w_{ij}\}_{i \neq j}$ held fixed — and may thus be maximized efficiently using any convex optimization method, e.g. Newton-Raphson or conjugate gradient ascent.

Our procedure therefore is to initialize the parameters for each neuron using some default values that we have found to be effective in practice in analyzing real data, and then iteratively (i) estimate the marginal posteriors via the SMC recursions (10-14) (E step), and (ii) maximize over the intrinsic parameters $\tilde{\theta}_i$ (M step), using the separable convex optimization approach described above. We iterate these two steps until the change in $\tilde{\theta}_i$ does not exceed some minimum threshold. We then use the marginal posteriors from the last iteration to seed the blockwise Gibbs sampling procedure described below for approximating $P[n_i, \mathbf{h}_i|\mathbf{F}; \theta_i]$.

2.4. *Estimating joint posteriors over weakly coupled neurons.* Now we turn to the key problem: constructing an estimate of the joint marginals $\{P[n_i(t), \mathbf{h}_i(t)|\mathbf{F}; \theta]\}_{i \leq N, t \leq T}$, which are the sufficient statistics for estimating the connectivity matrix $\mathbf{w}$ (recall Eq. (7)). The SMC method described in the preceding section only provides the marginal distribution over a single neuron's hidden variables; this method may in principle be extended to obtain the desired full posterior $P[\mathbf{X}(t), \mathbf{X}(t-\Delta)|\mathbf{F}; \theta]$, but SMC is fundamentally a sequential importance sampling method, and therefore scales poorly as the dimensionality of the hidden state $\mathbf{X}(t)$ increases (Bickel et al., 2008). Thus we need a different approach.

One very simple idea is to use a Gibbs sampler: sample sequentially from

$$(16)\quad X_i(t) \sim P[X_i(t)|\mathbf{X}_{\setminus i}, X_i(0), \ldots, X_i(t-\Delta), X_i(t+\Delta), \ldots, X_i(T), \mathbf{F}; \theta],$$

looping over all cells $i$ and all time bins $t$. Unfortunately, this approach is likely to mix poorly, due to the strong temporal dependence between $X_i(t)$ and $X_i(t+\Delta)$. Instead, we propose a blockwise Gibbs strategy, sampling one spike train as a block:

$$(17)\quad \mathbf{X}_i \sim P[\mathbf{X}_i|\mathbf{X}_{\setminus i}, \mathbf{F}; \theta].$$

If we can draw these blockwise samples $\mathbf{X}_i = \mathbf{X}_i(s : t)$ efficiently for a large subset of $t - s$ adjacent time-bins simultaneously, then we would expect the resulting Markov chain to

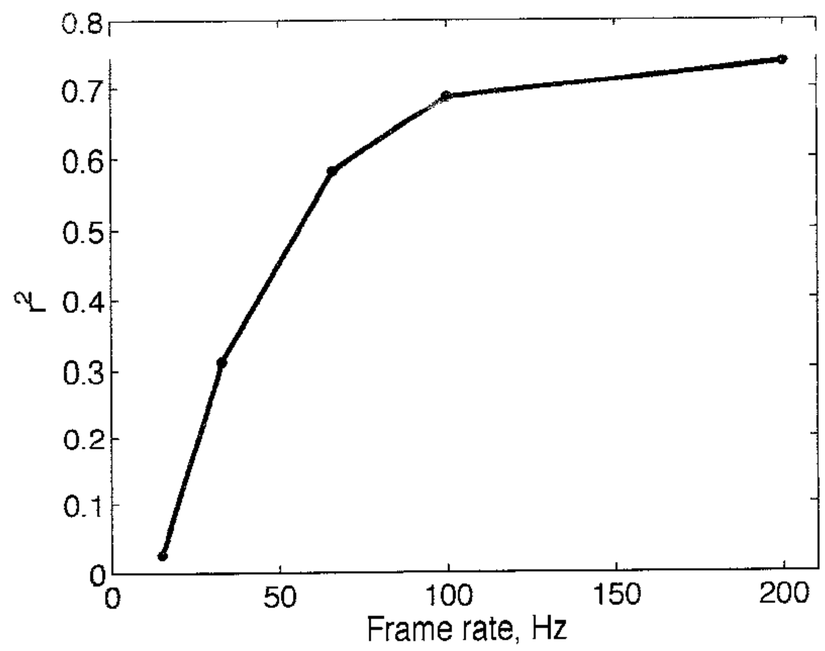


FIG 8. *Accuracy of the inferred connectivity as a function of the frame rate of calcium imaging. A population of $N = 25$ neurons firing at $\approx 5$ Hz and imaged for $T = 10$ min was simulated here, with eSNR $\approx 10$. At 100 Hz, $r^2$ saturated at the level $r^2 \approx 0.7$ achieved with $\Delta \to 0$.*

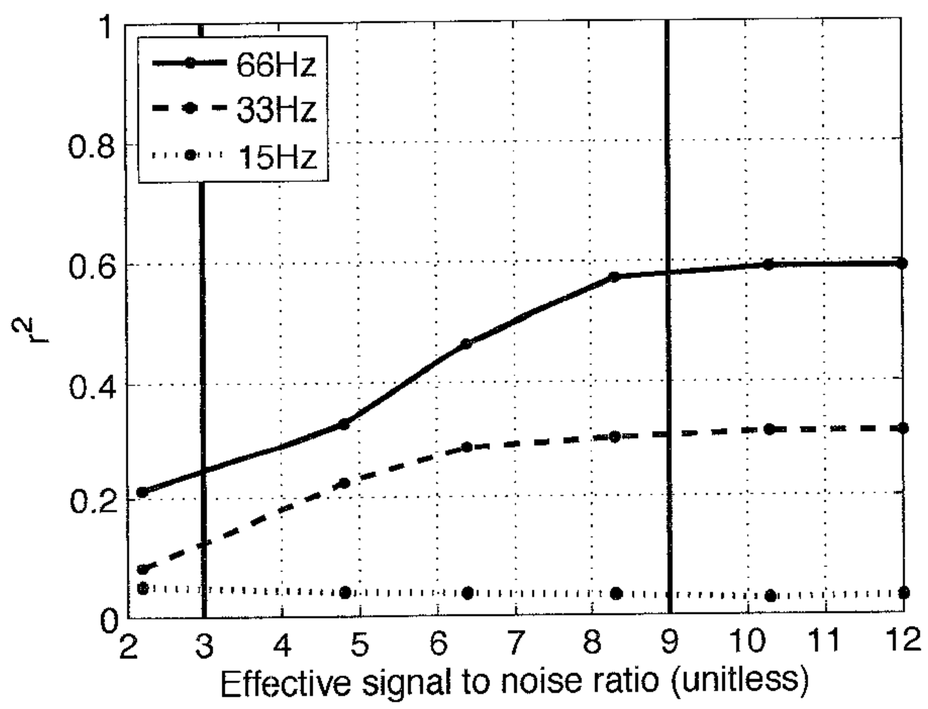


FIG 9. *Accuracy of inferred connectivity as a function of effective imaging SNR (eSNR, defined in Eq. 6), for frame rates of 15, 33, and 66 Hz. Neural population simulation was the same as in Figure 8. Vertical black lines correspond to the eSNR values of the two example traces in Figure 3, for comparison.*

neurons fired largely independently of each other: see Fig. 11, upper left for an illustration. In this case of weakly-correlated firing, the inputs $\{h_{ij}(t)\}$ will also be weakly correlated, and the model should be identifiable, as indeed we found. Should this weak-coupling condition be violated, however (e.g., due to high correlations in the spiking of a few neurons), we may require much more data to obtain accurate estimates due to multicollinearity problems.

To explore this issue, we carried out a simulation of a hypothetical strongly coupled neural network, where in addition to the physiologically-relevant weak sparse connectivity discussed in section 2.7 we introduced a sparse random strong connectivity component. More specifically, we allowed a fraction of neurons to couple strongly to the other neurons, making these "command" neurons which in turn could strongly drive the activity of the rest of the popula-